

**PCT**WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>6</sup> :</b> <b>C07H 21/04, C12P 19/34, C12Q 1/68</b>	<b>A1</b>	<b>(11) International Publication Number:</b> <b>WO 97/31012</b> <b>(43) International Publication Date:</b> 28 August 1997 (28.08.97)
<b>(21) International Application Number:</b> PCT/US97/02403 <b>(22) International Filing Date:</b> 18 February 1997 (18.02.97) <b>(30) Priority Data:</b> 60/012,061 22 February 1996 (22.02.96) US <b>(60) Parent Application or Grant</b> (63) Related by Continuation US 60/012,061 (CIP) Filed on 22 February 1996 (22.02.96) <b>(71) Applicants (for all designated States except US):</b> THE RE- GENTS OF THE UNIVERSITY OF MICHIGAN [US/US]; Management Technology Office, Wolverine Tower, Room 2071, 3003 South State Street, Ann Arbor, MI 48109-1280 (US). BOARD OF TRUSTEES OPERATING MICHIGAN STATE UNIVERSITY [US/US]; East Lansing, MI 48824 (US). <b>(72) Inventors; and</b> <b>(75) Inventors/Applicants (for US only):</b> BREWER, George, J. [US/US]; 3820 Gensley, Ann Arbor, MI 48103 (US). VENTA, Patrick, J. [US/US]; 9646 Rolling Green, Pinck-		ney, MI 48169 (US). YUZBASIYAN-GURKAN, Vilma [US/US]; 3101 Dexter Road, Ann Arbor, MI 48103 (US). <b>(74) Agents:</b> SMITH, DeAnn, F. et al.; Harness, Dickey & Pierce, P.L.C., P.O. Box 828, Bloomfield Hills, MI 48303 (US). <b>(81) Designated States:</b> AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).  <b>Published</b> <i>With international search report.</i>
<b>(54) Title:</b> GENE SPECIFIC UNIVERSAL MAMMALIAN SEQUENCE-TAGGED SITES  <b>(57) Abstract</b>  Primer sets which amplify conserved regions of specific genes across mammalian species are provided. The methods used to design the primer sets as well as methods of making and using the primer sets are also provided.		

**BEST AVAILABLE COPY**

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgyzstan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic	KR	Republic of Korea	SE	Sweden
CG	Congo	KZ	Kazakhstan	SG	Singapore
CH	Switzerland	LI	Liechtenstein	SI	Slovenia
CI	Côte d'Ivoire	LK	Sri Lanka	SK	Slovakia
CM	Cameroon	LR	Liberia	SN	Senegal
CN	China	LT	Lithuania	SZ	Swaziland
CS	Czechoslovakia	LU	Luxembourg	TD	Chad
CZ	Czech Republic	LV	Latvia	TG	Togo
DE	Germany	MC	Monaco	TJ	Tajikistan
DK	Denmark	MD	Republic of Moldova	TT	Trinidad and Tobago
EE	Estonia	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	UG	Uganda
FI	Finland	MN	Mongolia	US	United States of America
FR	France	MR	Mauritania	UZ	Uzbekistan
GA	Gabon			VN	Viet Nam

## GENE SPECIFIC UNIVERSAL MAMMALIAN SEQUENCE-TAGGED SITES

### FIELD OF THE INVENTION

The present invention relates generally to genetic markers and methods of  
5 making and using such markers, and more particularly, to primer sets referred to as  
universal mammalian sequence-tagged site primers, which may be used to amplify  
conserved regions of specific genes across mammalian species.

### BACKGROUND OF THE INVENTION

Over the last several years significant effort has been made to develop  
10 genome maps of many species. The subjects of the majority of these projects have  
been mammals, including human, mouse, rat, ox, sheep, pig, horse, cat, and dog  
(e.g., Buchanan F.C. et al., *Genomics* 22:397-403 (1994); Dietrich, W. et al.,  
*Genetics* 131:423-447 (1992); Ellegren, H. et al., *Anim. Genet.* 23:133-142 (1992);  
O'Brien, S.J., *Trends Genet.* 2:137-142 (1986); Serikawa, T. et al., *Genetics*  
15 131:701-721 (1992); Weissbach, J. et al., *Nature* 359:794-801 (1992); Wintero, A.K.  
et al., *Genomics* 12:281-288 (1991); and Barendse, W. et al., *Nat. Genet.* 6:227-235  
(1994)). For non-human mammals, genome maps will lead to important applications  
including optimized breeding strategies, wherein desirable characteristics are  
selected and unwanted genes are removed, particularly those genes that lead to  
20 genetic disease. Comparisons made between genome maps will also provide a  
wealth of scientific information (O'Brien, S.J. et al., *Nat. Genet.* 3:103-112 (1993)).

The traditional method for developing gene-specific markers, Southern  
blotting and cross-species hybridization, is time consuming, labor intensive and  
limited in flexibility. However despite its many limitations, this method remains the  
25 primary method for developing gene-specific markers in most mammalian genome  
projects. There thus remains a need for a more efficient method. A more efficient  
method is particularly important for non-human mammalian genome projects wherein  
scientific resources are often limited.

One potential method is the development of polymerase chain reaction (PCR)  
30 primers to conserved regions of specific genes that can be used across species.  
Although this method has been successfully used for the study of a number of  
individual genes, it has not been applied on a genome-wide basis for the purpose  
of map development.

It would thus be desirable to provide genetic markers which may be used to  
35 detect specific mammalian genes. It would further be desirable to provide genetic  
markers which may be used to develop mammalian genome maps. It would further

- 2 -

be desirable to provide genetic markers which may be used across mammalian species. It would also be desirable to provide a method for designing and producing genetic markers based on conserved regions of genes.

#### SUMMARY OF THE INVENTION

- 5       Primer sets which may be used to amplify conserved regions of specific genes, including eleven preferred primer sets that have been shown to amplify specific canine genes, are provided. The methods used to design the primer sets as well as methods of making and using the primer sets for example with a polymerase chain reaction (PCR), are also provided. Such genetic markers based
- 10   on PCR primers are often called sequence-tagged sites (STSs) or sequence-tagged site primers (Olsen, M. et al., *Science* 245:1434-1435 (1989)). Because the primer sets of the present invention may be used to locate genes across mammalian species, such primer sets are referred to herein as universal mammalian sequence-tagged site (UM-STs) primers.
- 15       Additional objects, advantages, and features of the present invention will become apparent from the following description and claims taken in conjunction with the accompanying drawings.

#### BRIEF DESCRIPTION OF THE DRAWINGS

- 20       The various advantages of the present invention will become apparent to one skilled in the art by reading the following specification and claims and by referencing the following drawings in which:

Figure 1 is a photograph of a gel showing the amplification of several canine gene segments using the UM-STs primers of the present invention;

- 25       Figure 2 shows the lineup of several canine gene sequences with homologous non-canine mammalian genes;

Figure 3 is a photograph of a gel showing the amplification of a portion of the *FES* proto-oncogene from several mammalian DNAs using the UM-STs primers of the present invention; and

- 30       Figure 4 shows sequence comparisons of a portion of the *FES* proto-oncogene from several mammalian DNA.

#### DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

The universal mammalian sequence-tagged site primer sets of the present invention are set forth in Tables 1 and 1A. The methods used to design the primer sets as well as methods of making and using the primer sets are also provided.

- 3 -

The primer sets of the present invention referred to herein as UM-STS primers, may be used to amplify mammalian genome regions of interest, isolate clones from mammalian genomic and cDNA libraries and perform cross-species genome comparisons. The UM-STS primers of the present invention will also be  
5 useful for developing additional genetic markers within various genomes. For example, a microsatellite repeat has been found within the retinoblastoma (*RB1*) loci disclosed herein. Single site variability may also be found directly in at least some of the amplified products by using one of a number of techniques developed for scanning for variability, such as the single-strand conformation polymorphism  
10 technique. For example, this method has been used to find two polymorphic sites in a study of the canine *ALAS2* gene in a PCR product of a size similar to those set forth herein (Boyer, G. et al., *Anim. Genet.* 26:206-207 (1995)). If the frequency of single site polymorphic variability for other mammals is as high as that estimated for humans (roughly one in 200 to 400 nucleotides), then a significant portion of UM-  
15 STSs will have these sites. Each species will be screened individually for genetic variability using UM-STS primers.

An example of the usefulness of cross-species genome comparisons is given by the case of Waardenburg syndrome. Direction to the location of one of the human Waardenburg syndrome genes, well known for causing a syndromic hearing  
20 loss, was first found from comparative mapping with the mouse (Asher, J.H.J. et al., *J. Med. Genet.* 27:618-626 (1991)). The map locations in the mouse suggested possible locations of the human disease gene, one of which eventually was proven correct (Morrell, R. et al., *Hum. Mol. Genet.* 156:53-59 (1993)). Because the identity of the gene in the mouse was not known at the time, this approach might  
25 more properly be called a 'candidate linkage' approach. UM-STSs will thus be useful not only for rapidly producing genetic maps but the candidate linkage approach may also be applied to more species.

Very little is known about the location of genes within the canine genome. The development of UM-STSs will help to rapidly identify the location of linkage  
30 groups on specific canine chromosomes. The assignment of conserved syntenies will allow candidate linkages to be tested in the canine genome. The assignment of the proposed anchor loci (O'Brien, S.J. et al., *Nat. Genet.* 3:103-112 (1993)) as defined by UM-STSs to specific chromosomes may be accomplished by the somatic cell hybrid, flow sorted chromosome, and fluorescent in situ hybridization (FISH)

- 4 -

methodologies. Other methods known to those skilled in the art, such as assignment by use of linkage to previously mapped loci, are also possible.

### SPECIFIC EXAMPLE 1

The following example further describes the UM-STS primers of present invention as well as methods of designing, making and using such primers.

#### Materials and Methods

*DNA Isolation* - DNA from dog, human, pigtail macaque, horse, pig, rat and mouse were isolated from various tissues by standard phenol-chloroform extraction methods (Sambrook, J. et al., *Molecular Cloning. A Laboratory Manual*. (2nd ed.) Cold Springs Harbor: Cold Springs Harbor Laboratory Press (1989)). Goat DNA was supplied by Dr. Karen Friderici, Michigan State University. DNA was purified by standard methods from a canine liver cDNA library (Clontech) and from a canine genomic DNA library (Clontech) after growing  $1 \times 10^6$  phage in *E. coli* strain LE392 (Murray, N.E. et al., *Mol. Gen. Genet.* 156:53-59 (1977)) in liquid culture (Sambrook, J. et al., *Molecular Cloning. A Laboratory Manual*. (2nd ed.) Cold Springs Harbor: Cold Springs Harbor Laboratory Press (1989)).

*Design of PCR Primers* - Primers were designed to genes where the intron-exon structure was known in at least one species and where the nucleotide sequence was known in at least two species (the index species) that were not closely related. Tandemly duplicated genes known to have undergone gene conversion in any species were avoided. Primers were generally designed so that the amplified product contained an intron. The human gene nomenclature system (ISGN, *Cytogenet. Cell Genet.* 46:11-28 (1987)) was followed for naming the canine genes. An example of the loci used and their protein products are: *CFTR*, cystic fibrosis transmembrane regulator; *COL10A1*, type X collagen, alpha 1 chain; *CSF1R*, colony stimulating factor 1 receptor; *CYP1A1*, cytochrome P-450 1, alpha 1; *DCN1*, decorin; *FES*, c-fes (feline sarcoma) proto-oncogene; *GHR*, growth hormone receptor; *GLB1*, beta galactosidase; *PKLR*, pyruvate kinase - liver, RBC form; *PVALB*, parvalbumin; and *RB1*, retinoblastoma protein. The Genbank Accession numbers or reference for the sequence of the two index species for each locus are as follows: *CFTR*, M55129, M60493; *COL10A1*, X65120, X65121; *CSF1R*, X14720, K01643; *CYP1A1*, (Uchida, T. et al., *Mol. Pharmacol.* 38:644-351 (1990)), X04300; *DCN1*, L01125, Z12298; *FES*, X06292, J02088; *GHR*, Z11802, J04811; *GLB1*, S59584, M57734; *PVALB*, X63578, M15452; *PKLR*, S59798, M17088; and *RB1*, L11910, M26391.

- 5 -

Primers were designed to highly conserved nucleotide sequences contained within coding regions. Additional considerations taken into account were: degeneracy of underlying codons (Li, W.-H. et al., *Fundamentals of Molecular Evolution*. Sunderland, MA: Sinauer Associates, Inc. (1987)), overall amino acid mutability of the primer region (Collins, D.W. et al., *Genomics* 20:386-396 (1994)), placement of the 3' end of the primer with respect to amino acid mutability, as well as standard design practices for primers known in the art such as avoidance of primer-dimers. Conservation of amino acids within multigene families was also taken into account, when possible. Where unavoidable mismatches occurred between the two index species, the primer sequence was designed to match one of the two which was then designated the primary index species. GC-rich genes were avoided due to amplification difficulties, even with exactly matching primers. Primers were twenty bp in length on average. Each primer in a pair was adjusted to be of approximately the same annealing temperature (Breslauer, K. et al., *PNAS (USA)* 83:3746-3750 (1986)). All sets of primer pairs were designed to have approximately the same annealing temperature in anticipation of performing multiplex amplifications. It was not always possible to follow every rule for every gene, given the actual circumstances; however, the majority of the rules were generally applicable.

*PCR Amplifications* - Correct design and syntheses of the primers were examined by amplifying the DNA from the primary index species. Standard buffer, nucleotide, and primer concentrations were 50 mM Tris-HCl (pH 8.3 at room temperature), 50 mM KCl, 1.5 mM MgCl<sub>2</sub>, 200  $\mu$ M dNTPs, 0.1  $\mu$ g of each primer, and 0.5 - 1.0  $\mu$ g of target DNA in a 25  $\mu$ l reaction. Reactions were routinely boiled for three min prior to the addition of 2.0 U of Taq DNA polymerase. Optimal cycling conditions for the amplification of canine genomic DNA were usually found by testing one of several sets of conditions in general use in the lab (see Table 2). Occasionally it was necessary to use "hot-start" conditions (Bassam, B.J. et al., *Biotechniques* 14:30-34 (1993)) in order to get stronger, cleaner amplifications. The presence of an amplification product was determined by electrophoresis of a portion of the reaction on a 1% agarose 1X TBE gel (90 mM Tris, pH 8.3, 90 mM sodium borate, 2.5 mM EDTA) followed by staining with ethidium bromide.

*DNA Sequence Analysis* - The identity of each amplified canine gene was confirmed by 'single pass' direct sequencing of PCR products using Sequence or Taq DNA polymerase (United States Biochemical Corp., Cleveland). The PCR products were gel purified with Qiaex (Qiagen Corp., Chatsworth, CA) or by elution

- 6 -

from polyacrylamide gel slices (Bergenheim, N.C.H. et al., *Biochem. Genet.* 30:279-287 (1990)) prior to their use in the sequencing reactions. The canine sequences were aligned visually with the sequences of the other species used to design the PCR primers in order to verify the degree of sequence identity.

5

## Results

The primer sets for the various UM-STSS reported here are given in Tables 1, 1A and 2 and efficient amplification conditions for the eleven preferred UM-STSS are given in Table 3. It will be appreciated by those skilled in the art that the conditions may be optimized further (e.g., reduction in the time in each cycle).

10 However, the conditions set forth herein were found to work effectively while minimizing the number of conditions that had to be examined.

A representative gel showing amplification of the canine target DNA along with the human target DNA is shown in Figure 1. In Figure 1 the various lanes were amplified with the following gene-specific primer sets of the present invention: lanes 15 1-4, *GHR*, lanes 5-8, *COL10A1*, and lanes 9-12, *DCN1*. Lane 13 contained a mixture of DNA size markers; lambda bacteriophage DNA cut with the restriction endonuclease BstE II and the plasmid pSK- (Stratagene) cut with Msp I. Lanes 1, 5, and 9 contain PCR products amplified from human genomic DNA. Lanes 2, 6, and 10 contain PCR products amplified from canine genomic DNA. Lanes 3, 7, and 20 11 contain PCR products amplified from DNA purified from a canine genomic library contained in a lambda phage vector. Lanes 4, 8, and 12 contain PCR products amplified from a canine liver cDNA library. The human target serves as a positive control for the amplification system because these primers were designed to exactly match the human sequence. The ability to quickly screen genomic and cDNA 25 libraries for the presence of sequences is also demonstrated in Figure 1. The genomic clones for *GHR*, *COL10A1*, and *DCN1* (a very faint signal, stronger on other gels [data not shown]) are present in this particular canine genomic library. The presence of a decorin cDNA clone (encoded by the *DCN1* locus) in the canine liver cDNA library is shown by the presence of the 122 bp band; cDNA clones for 30 *GHR* and *COL10A1* are not present. The *DCN1* PCR product from the cDNA library was sequenced and its identity confirmed (see Figure 2). The human and canine genomic bands have different sizes for *GHR* and *DCN1* because of the intron size differences. The size for the *COL10A1* PCR product is the same between the species because this the only preferred UM-STSS set forth herein in which an intron 35 was not spanned. Although the PCR product bands in Figure 1 are unique, some



- 7 -

UM-STS-species combinations contained one to several non-specific amplification products. However, it is almost always possible to deduce the correct band based upon staining intensity and the similarity in size compared to the band of the primary index species.

5       The amplified products for all of the canine loci (see Table 2) were sequenced to confirm their identity and the results are shown in Figure 2. The locations of PCR primers are underlined, although not all PCR primer sites are shown. Some of the lineups show intron sequence whereas others simply identify the location of the introns. The degree of identity between the canine and index  
10 species sequences for each locus is within the range generally accepted (approximately 70 to 100%) as indicating homology for mammalian species (Li, W.-H. et al., *Fundamentals of Molecular Evolution*. Sunderland, MA: Sinauer Associates, Inc. (1987)). These results support the hypotheses that the canine PCR products are homologous to the respective index species' genes. The canine  
15 *COL10A1* sequence matched the human and mouse sequences to a similar extent (data not shown). The sequences for *PKLR* and *CYP1A1* exactly matched previously published canine coding sequences (Whitney, K.M. et al., *Exp. Hematol.* 22:866-874 (1994); and Uchida, T. et al., *Mol. Pharmacol.* 38:644-351 (1990)); the sequence for canine *FES* is shown in Figure 4. Although the majority of the canine  
20 sequence for *PVALB* is from an intron, it is believed that the degree of sequence identity from this region is sufficient evidence to confirm that the PCR product is from the correct canine locus. As expected, the canine sequences tend to show greater identity with the human sequences than with the rodent sequences because of the faster evolutionary rate of the rodent genome (Gu, X. et al., *Mol. Phylogen. Evol.*  
25 1:211-214 (1993)). As set forth above, a microsatellite repeat was found within the amplified product itself for *RB1*. Preliminary results show that the *RB1* repeat, (GA)<sub>12(avg)</sub>, has moderate genetic variability within several canine breeds.

The 'universal' utility of these primers was studied on the DNAs from mammals representing several different orders using the primer sets under the  
30 reaction conditions that were found to amplify the canine sequences. These reactions are referred to herein as 'Zoo PCRs.' Figure 3 shows a representative experiment. The *FES* proto-oncogene amplified from all of the mammalian species were examined. These DNAs were purified and sequenced directly without subcloning (see Methods and Materials above). In Figure 3, target DNAs for each  
35 lane are: 1, human; 2, pigtailed macaque; 3, dog; 4, goat; 5, pig; 6, horse; 7, mouse;

- 8 -

and 8, rat. The mouse DNA was degraded; strong amplification was obtained with another lot (sequence shown in Figure 4). The DNA marker lane (M) contains a 100 bp ladder. The sequences are tabulated in Figure 4. Sequences are from exon 15 and intron 15. Notations for the sequence lineups in Figure 4 are HUM, human; 5 MAC, macaque; CAT, domestic cat; FES, feline sarcoma virus; DOG, dog; COW, ox; GOA, goat; HOR, horse; PIG, pig; RAT, rat; MOU, mouse. The upper two lines for each block of text represent amino acid sequences and the lower lines represent nucleotide sequences. Dots indicate nucleotides in the various species that are identical with that of the human sequence. The human and cat sequences match 10 the published sequences (Alcalay, M. et al., *Oncogene* 5:267-275 (1990); and Roebroek, A.J.M. et al., *J. Virol.* 61:2009-2016 (1987)). The feline sarcoma virus sequence was not determined but is included for comparative purposes. Only a single amino acid interchange was found among these sequences; isoleucine (I) for macaque, cat, and feline sarcoma virus and leucine (L) in all others. Sequence 15 alignments for the intron were done visually.

The degree of sequence identity makes it highly likely that the canine PCR products are homologous with the corresponding index species' genes. The pattern of nucleotide interchange is also what would be expected for homologous genes; members of the same mammalian order share more sequence similarity with one 20 another than with those of other orders.

The data for the Zoo PCRs for the preferred UM-STS primer sets are given in Table 4. Greater than eighty-four percent of the targets, excluding the index and canine species, amplified under the single condition used to amplify the canine sequence. These species represent five different mammalian orders; primates 25 (human and macaque), carnivores (dog), artiodactyls (goat and pig), perissodactyls (horse), and rodents (mouse and rat). Limited experiments on other members of these orders (e.g., cat and ox) produced similar results (data not shown). Lack of amplification for *DCN1* for one of the artiodactyls (goat) would be predicted because there are four mismatches between the UM-STS primers and the sequence 30 of the closely related bovine *DCN1* (Day, A.A. et al., *Biochem. J.* 248:801-805 (1987)).

**Table 1**  
**Primer sequences for UM-STs sorted according to chromosomal location.**

Seq'd Locus Name	Gene Product	Human Chromosome Band	Primer 1 Sequence	Primer 2 Sequence
NRAS	N-RAS Proto-Oncogene	1 p13	ATGACTGAGTACAAACTGGT	CTTCGGCTGTCCTCATGTA
X PND	Pronatriodiatin	1 p36	GCAGACCTGCTGGATTTCAG	CAGTCGGCTCTGGGCTCCAAT
FGR	FGR Protooncogene	1 p36.1	GAGTGGTACTTTGGAAGATTGG	TAGTCTGCACACAGCTCCTG
D CXN40	Connexin 40	1 pter-q12	ATGGGTGACTGGAGCTTCCT	CCCACAATGAAGGCCACCTC
X PKLR	Pyruvate Kinase - RBC	1 q21	CGCCTCAAGGAGATGATCAA	ATGAGCCCGTCGTCAATGTA
X AT3	Anthrithrombin III	1 q23-q25	CTTCCTTGGCCAAACTGAAC TG	GGGCTGAACCTTGACTTCCA
X REN	Renin	1 q32	ACACTCCCGACATCTCTTT	CGCCGATCAAACTCTGTA
X SFTF3	Pulmonary Surfactant Protein 3	2 p11.2	GGAAGTTCTCTGGAGCATGAG	CACAGGCCCCAGGTGCTTACA
X SPTBN1	Beta Spectrin (Non-RBC)	2 p21	TCTCAAGACTATGGCAAACA	CTGCCATCTCCCAGAAGAA
X APOB	Apolipoprotein B	2 p24-p23	TATGATTTTCTCTTAATCAAG	GTGCCCTCTAATTTGTACTG
X IL1A	Interleukin 1 Alpha	2 q13	AGAACTCAAGATGGCCAAAGT	TGATTCAGAGACAGATGGTC
X COL3A1	Collagen III Alpha 1	2 q31-q32.3	GGACCAAGGAGTATGGGAA	ACTTTCTCTTGACTTCCCT
X ELN	Elastin	2 q31-qter	GGTGACGCGCTAAAGCAG	AGGACACCTCCAAGGCCAG
X PAX3	Human Paired Domain 2	2 q34-q36	GCCACAAGATCGTGGAGATG	GGTTCTCTCTTTTGTATTCTC
X GCG	Glucagon	2 q36-q37	TTCATTGCTTGGCTGGTGAA	GTGTTCACTCTCATCAGAGAA
X PIT1	Pituitary-Specific Transcription Factor 1	3 p11	TTCAGTCAAAACAACAATCTG	GCTCCCACITTTTTCATTGTA
X GNAT-1	Rod Transducin Alpha	3 p22-21.3	CGCAAGAAGTGGATCCACTG	GCAGATGCTGTTGAACAGGT
RAF1	c-RAF Protooncogene	3 p25	CACATCAACAACCCGAGATCA	GCATTGATATCCTCAGTGTG
X GLB1	B-Galactosidase	3 pter-p21	GAATTCATACTGGCTGGCT	CATCCCAATAGGCCAAAATTGGT
X GPX1	Glutathione Peroxidase 1	3 q11-q12	GACTACACCCAGATGAACGA	CAGGAACCTCTCAAAGTTCC
X TF	Transferrin	3 q21	GCTGACAGGGACCAGTATGA	AACAGCAGGTCTCTCCCATG
X RHO1	Rhodopsin	3 q21-qter	TACATGTTCTGGTCCACTT	TGGTGGGTGAAGATGTAGAA
X GLUT2	Glucose Transport-like 2	3 q26.1-q26.3	TGGATGAGTTATGTAGCAT	GACTTTCCTTTGGTTTCTGG
X SST	Somatostatin	3 q28	GACTCCCGAGGCTTCTCTTTG	ATACTGCAGGAGAGAGAAGAA
X HOX7	Homeobox 7	4 p16.1	AAGTCCGCCAGAAAGCAGTA	ATCTTCAGCTTCTCCAGCTC
X PDEB	cGMP Phosphodiesterase Beta	4 pter	CTGAAGAGCTACTACACGGA	TGACACTGTTTCATCCACCA
X ALB	Albumin	4 q11-q13	GGCTGACTGCTGTGCAAAACA	AAGTAAGGATGCTCTCTGGC
X KIT	c-KIT Protooncogene	4 q12-q13	CCTGTGAAGTGGATGGCACC	GCATCCCAGCAAGTCTTCAT

-10-

Table 1 (cont.)

Seq'd Locus Name	Gene Product	Human Chromosome Band	Primer 1 Sequence	Primer 2 Sequence
X FGG	Fibrinogen Gamma	4 q28	CAATATAAGAAGGATTGGACA	TGACACTTGTTTCATCCACCA
X F11	Coagulation Factor XI	4 qter	AATCTTTGTCCTTTAAACATC	GGGTATAGGTAAAAACTGGCA
X GHR	Growth Hormone Receptor	5 p13.1-p12	CCAGTTCAGTTCCTCAAGATC	TGATTCCTCTGGTCAAGGCA
X HEXB	Beta Hexosaminidase	5 q13	TTTCATGGTGGAGAAGCTTG	ATCTTTGGAACCTCCAGAGTC
X HMGCR	HMG CoA Reductase	5 q13.3-q14	GGATATATGCCATCCCTGT	GTCTCCCATCTGCAAGGAC
X IL4	Interleukin 4	5 q23-q31	CTATTAAATGGGTCTACCTCCAACT	TCAACTCGGTGCACAGAGTCTGG
X PDEA	Rod Phosphodiesterase Alpha	5 q31-33	TGCCAAACAGTACTACCAACC	TGCAGCTCCAGGTGGTTGA
X ADRB2	Adrenergic Receptor Beta 2	5 q31-q32	CCATTTCAGATGCACTGGTA	GCAGCCAGCAGAGGGTGAA
X CSF1R	CSF-1 Receptor	5 q33-q35	TTCCAAACACGGGGACCTA	CATGCCAGGGCGGAGAGGA
X TNFA	Tumor Necrosis Factor Alpha	6 p21.3	CTCAGGCTCTTCTCCTTCTCT	ATGGGCTCATACCAAGGCTT
X EDN1	Endothelin 1	6 p24-p23	CCAAAAGACAAGAGTGCTG	TGGAACAGTCTTTTCTCTTCT
X COL9A1	Collagen IX Alpha 1	6 q12-q14	ATCAGGATTGGCCCAAGATGA	GGAACTCTGAAGTCTACATT
X SOD2	Superoxide Dismutase 2	6 q21	AAGTTTAAAGAGAAGCTGAC	TCCAGTTGATTACATTCCA
X COL10A1	Collagen Type X Alpha 1	6 q21-q22	ATTCTCTCCAAAGCTTACCC	GGCACTAGGAATCCTGAGAA
X PLG	Plasminogen	6 q25-q27	CAGTCCCTGTGATTGAGAA	TAGACACCAAGGCTTATTGGG
X ASL	Argininosuccinate Lyase	7 cen-q11.2	ATGACCATCTCAGCAGGAT	AAGTCCCTTGAGGGTCAATCA
X EPO	Erythropoietin	7 q21	CTCCCTCTGGGCTCCCACT	CCATCCTCTCCAGGCATAGAA
X CFTR	Cystic Fibrosis Trans. Regulator	7 q31-q32	CTAAGCCATGGCCACAAAGCA	CATTGCTTATCCTCTGTGT
X TCRB	T-Cell Receptor Beta	7 q35	GACTGTGGCTTCACCTCGG	GATCTCATAGAGGATGGTGG
X SFTP2	Pulmonary Surfactant Protein 2	8 p21	CAGAACACACAGGAGATGGT	GCCATCTTCATGATGTAGCA
X CA2	Carbonic Anhydrase II	8 q22	CAGTCCATTTTCAGTGGGG	GGCCAGTCCATCAGGTTGCT
X TG	Thyroglobulin	8 q24	TTCACTCAGAGTGCTACTG	GCCTTCTGTAGCTCATGATCTT
X GGTB2	Beta-1,4-Galactosyltransferase	9 p13	GGGAGGTGCGCATGATCC	ATTGGCCACTGTGATTGGGT
X ALDOB	Aldolase B	9 q21.3-q23	GTGACTGCTGGACATGCCTG	TTTGACGCTTGCCACCCC
X C5	Complement Factor 5	9 q22-q34	TGTGTACGATCCGGATATTGA	GCTCCTTCACAGACTTTCTG
X ABL	ABL Proto-oncogene	9 q34	GAGGACACCATGGAGGTGGA	GTGGATGAAGTTCTTCTTCTC
X IL2R	Interleukin 2 Receptor	10 p15-p14	CCCTTCCACATGGATTGAAA	CATCCAGTTAGCATATACAC
X RET	RET Proto-Oncogene	10 q11.2	GCTCTACAGAGAGTCCCTG	GGCAGGAAGTCTCACTCTC
X PSAP	Prosaposin	10 q21-q22	AAGTGTGGAAGAAGCTGGT	AAGGAAGGATCCATCACCTC
X TDT	Terminal Transferase	10 q23-q24	ACCTGGAAAGGCCATCCGTG	CGCCGAGGGTCTCTCTCAAA
X OAT	Omithine Aminotransferase	10 q26	CGTGTCTTCAGGATCCAA	GCCAGCCATCTACCAGTTCT
X WTI	Wilms Tumor 1	11 p13	GAGAAACCATACAGGTGTA	GTTTACCTGTATGAGTCTCT
X LDHA	Lactate Dehydrogenase A	11 p14-15.5	AACCTCAAGCTGGTCAATT	GAATCCAGATTGCAACCACT
D INS	Insulin	11 p15.5	GAGCCGGCTTCTTCTACAC	GGTAGAGGGAGCAGATGCTGG
X CD20	CD20	11 q12-q13.1	CTCTTGTGCTGCTTCTGGAAT	TGGAAAGAGGCAAGATCAGCAT
X ROM1	Rod Outer Segment Protein-1	11 q13	CAGAGGACGGCCACAGAA	GTTAAACACCAACAGAGGCTT
X APOC3	Apolipoprotein C3	11 q23-qter	CAGGAACAGAGGTGCCATGTC	TGCGCCACCTGGGACTCCTG

Table 1 (cont.)

Seq'd Locus Name	Gene Product	Human Chromosome Band	Primer 1 Sequence	Primer 2 Sequence
X VWF	von Willebrand's Factor	12 p	CCAGAGCGCATGGAGGCTG	CTGCACCTCCAGCTTGAATCC
X LDB	Lactate Dehydrogenase B	12 p12.1-12.2	TTCTCAGATCGTCAAGTACA	CTGCTGGGATGAATGCCAAG
X IL6	Interleukin 6	12 p12.2-p12	GCACTGGCAGAAACCAACCT	ATCTGAACCTCCACAAGACC
X TPI	Triphosphate Isomerase	12 p13	TATATGACTTCGCCGGCA	ATGCCACACAGGCTCATA
X COL2A1	Collagen II Alpha 1	12 q14.3	CTCTGCAGCAGACATAATCTG	TCTCCAGGTTCTCCTTTCTG
X DCN1	Decorin	12 q21-q23	GTTGATGCAGTAGCCTGAA	AAGTGAAGTCCCTCAGATG
X IGF1	Insulin-Like Growth Factor 1	12 q22	GGCATGTGGATGAGTGTG	CTCCTCTGTTCCTCCTCTG
X PLA2	Phospholipase A2	12 q23-q1er	GACTACGGCTGCTACTGTG	TTACAGCTGGCCAGTTCTT
X RB1	Retinoblastoma 1	13 q14.2	GTTCAGAAATAATACAGATGT	ACTCATTTCTGCCAGTTTCTG
X F7	Clotting Factor VII	13 q34	AATGGAGCTCAGTTGTGTG	CGATGTGTTGGTTGGTGGT
X CHY	Chymase (Mast Cell)	14 q11.2	GTCCACCTGGGAGATGTG	TGGGAGATTCCGGGTGAAGAC
X FOS	FOS Proto-Oncogene	14 q21.3	TCTGTGGCCCATCGCAGAC	ATCTTGACGGCAGGTGCGGTG
X CKBB	Creatine Kinase Brain	14 q32.3	TGGATCAAGGAGGAACCA	TTCAACCATCCACCACCAT
X TCRA	T-Cell Receptor Alpha	14 q34	ACTGTCTGCCCTGTTCCACGATT	GTAACAACTTGGCATCACAGGAAT
X B2M	Beta-2 Microglobulin	15 q21-q22.2	TTCCAGCAAGGACTGGTCTTT	CTGCTTACATGTCTCGATCT
X CYP1A1	Cytochrome P-450 (AH-I)	15 q22-q24	TTGGACCTCTTTGGAGCTGG	TGGTTGATCTGCCACTGGT
X PKM	Pyruvate Kinase - Muscle	15 q22-q1er	GCCTTCATTCAGACCCAGCA	ATTCCAGACTTAATCATCTCCTT
X FES	FES Proto-Oncogene	15 q25-q1er	GGGGAACCTTTGGGAAAGTGT	TCCATGACGATAGATGGG
X HGBA	Alpha Hemoglobin	16 p13.3	CCCACCAACCAAGCCTACTT	CGGTATTTGGAGGTCAGCAC
X GOT2	Glutamate Oxaloacetate Transaminase 2	16 q21-q22	TTTAAGTTACGCCGAGATGT	CTTGTAGGCCATGTCAAA
X CTRB	Chymotrypsinogen	16 q22.3-q23.2	AACGACATCACCTCTGCTTT	TGCAGGAGGAGAGCGCCACT
X APRT	Adenosine PR Transferase	16 q24	GACTCCCGAGGCTTCTCTTTG	ATACTGCAGGAGAGAGAAGAA
X TP53	Tumor Protein 53	17 p13.1	TACAAGCAGTCACAGCAGAT	TCTTCCAGTGTGATGATGGT
X GUC2D	Guanylate Cyclase 2D	17 p13.1	AGCTTACCGCATCCACGTGA	TTTGGGGATGGGCTTGTGA
X NF1	Neurofibromatosis 1	17 q11.2	ATTCACTCTCTGTGTACTTG	CAAAGCTCTGTGACTGTTT
X KRT10	Keratin 10	17 q12-q21	AGTACCCGACGCTGTAGA	TGACCTTGGTCCCTTAGATGA
X PDEG	Rod Phosphodiesterase Gamma	17 q21.1	AGGAAGCACAGCTTCTCAT	GACATCCCTCGGAATGGAAG
X SCN4A	Skeletal Muscle Sodium Channel	17 q23.1-q25.3	CTCAAGGTGGACATCTGTACAA	AGCAGCTCCGAGATGCCCTT
X TS	Thymidylate Synthetase	18 p1er-q12	TGCCAGTTCTATGTGTGAA	AGGTAAATATGTGCATCTCC
X TTR	Transferrin	18 q11.2-q12.1	GACTGGTATTTGTGTGAGG	CAGAGCAAAATGGCTCCCA
X C3	Complement Factor 3	19 p13.3-p13.2	GGAGTGGACTATGTGTAAGAC	TCAGGCCAGTGTCCACCCCT
X APOC2	Apolipoprotein C2	19 q13.2	GAATCACTCTACAGTTACTGG	AGCTGCTGTCTTTTGTCTGA
X CKMM	Creatine Kinase Muscle	19 q13.2-q13.3	AAGAAGCTCCGGGACAAAGGA	CAGCCACCGGTTCATGATGAA
X MX1	MX Anti-Viral Protein	21 q22.3	TGAAGAAGGGTTACATGATTGT	GGCAGGAACCGTGGCCTT
X PVALB	Parvalbumin	22 q12-q13.1	ATGTGAAGAAGGTGTTTCACAT	TCCTTGTCTCCAGCAGCCAT
X MAOA	Monamine Oxidase A	X p11.4-p11.23	AGGAACGGAAAGTTTGATAG	CTCCTTGTAAATACATCATGC
X DYS	Dystrophin	X p21	GTTTCAGGCCAGACCTCTTT	TACCGACCTTCAGGATCAAG

Table 1 (cont.)

Seq'd Locus Name	Gene Product	Human Chromosome Band	Primer 1 Sequence	Primer 2 Sequence
X MNK	Menkes Protein	X q12-q13.3	GGCATGACTTGTAAATTCCTG	CATCAAATCCCATGTCTTCTAT
PGK1	Phosphoglucokinase	X q13	ATTACCTTGCCTGTTGACTT	AAAGCTTCCCATTCAAATAG
PLP	Proteolipid Protein	X q21.3-q22.3	TTAGAGTGCTGTGCAAGATG	CCTAGCCATTTTCCCAACA
X HPRT	Hypoxanthine PR Transferase	X q26	AGCTTGCTGGTGAAAGGAC	TTATAGTCAAAGGGCATAATCC
X F9	Clotting Factor IX	X q26.3-q27.1	TGGGTGGTAACTGCAGCCCACT	CTACGCACACTCTTCACCCCA
F8	Clotting Factor VIII	X q28	GATGCACAGATTACTGCTTC	GTAAGCAGAGATTTTACTCCCTG
X SRY	Sex Determining Region - Y	Y p11.3	AAGCGACCCCATGAACGCATT	TCGGGTATTTCTCTCTGTG

Note: X in Seq'd (sequenced) column indicates PCR product sequenced;

D indicates primers designed to canine sequence from literature references.

Table 1A  
Primer sequences for UM-STs sorted according to chromosomal location.

S <sup>1</sup> Locus Name	Gene Product	Human C <sup>1</sup> some	Band	Primer 1 Sequence	Primer 2 Sequence
VCAM1	Vascular CAM 1		1 p32-p31	CCAGAGAAACCAATTACTGTTG	CCCTTAAGTAATTCATCTCTCT
GNB1	Guanine-Nucleo. Bind. Protein - Beta		1 pter-p31.2	ATGAGTGAGCTTGACCACT	GTTTGTGATCTGAGAGAGAG
CD23	CD23		1 q21-q23	TTCCAACGGGAAGTGCTACTA	GCATCATCAGCAGTCTCTC
AGT	Angiotensinogen		1 q40-q42	GCAGGACCTGCTCGCCAG	TGTTGGGTAGACTCTGTGGGCT
X SAG	S-antigen		2 q37	AGTGATGTCACCTCTGACC	TGTTGCTCCCAAGCTTCTT
X CNCG1	CGMP-Regulated Channel		4 p14-q13	GACCCAGAGAAAGAAAGAA	TACCCAGGAAGCAATCACA
EGF	Epidermal Growth Factor		4 q25-q27	GTGCTTGAGAGAGATGGAA	AAGACATTTTCGATAGCGGA
GRL	Glucocorticoid Receptor		5 q31-q32	CCCAAACTCTCGCTGGTGTG	TTGCTATCCACTGAATCTC
X RDS	Peripherin/RDS		6 p	TTTGACCAAGAAAGCGGGT	AGGAGCATCTGGTCCCTCTG
X GCAP	Guanylate Cyclase Activating Protein		6 p21.1	GATTGACGTCAACGGGGATG	GTCATGGCGTCTCGGGGAT
GCK	Glucokinase		7 p15 (p13)	GAGCAGATCCTGGCAGAGTT	TGAAATCACACAGCCCTCTT
X NGT1	Transducin-gamma		7 q21.3	GTGTCAAATGTTGTGAAGA	GAGAACTACAGGCGCTGTG
PLAT	Plasminogen Activator - Tissue		8 p12-q11.2	AAGCAACCGGGTGGATATTG	CGCATCTATCGGGGCATG
MTS1	Multiple Tumor Suppressor 1		9 p21-22	CAGGTGATGATGATGGCCAG	GTAATAGTGTGTCCAAACAA
X TYRP1	Tyrosinase Related Protein 1		9 p23	GCAAAGCGCACTCACTCC	CATGTGAGAGAAATGGCTG
ALDH1	Aldehyde Dehydrogenase 1		9 q21	AAACAGCAGAGCAAACTCC	CATTGAGCAGGCTTGGTATTG
VIM	Vimentin		10 p13	GCTGCCAACCGGAACAATGA	TTGTGCTGGAGCATCTCT
IRBP	Interstitial Retinal Binding Protein		10 q11.2	GAACGTATGGCTCCCAAGAA	GGCAGGCAGATGGTCTGTAT
PLAU	Plasminogen Activator - Urokinase		10 q24-qter	ACCACCATCGAAGAACCGCC	CCCTCTCTCTGCAATCCCA
GOT1	Glutamate Oxaloacetate Transam. 1		10 q24.1-q25.1	GCACCTCGAATTGGAGCTGA	GCTCTGATACAGCTGTGGT
TYR	Tyrosinase		11 q14-21	GAACAAGCGAGTCGGATCTG	TCCTCAGCATGAAGTCTGGAT
IGHF	Immunoglobulin E		14 q32.33	ACCTGCTGTGGTGGACCT	CCTCGGGTTCACCCCTCAGT
X AGC1	Aggrexin		15 q25-q26.2	ATGGACATGTCAGTCCGGCT	GCTTGAGAAGGCCTTGTAACC
IL4R	Interleukin 4 Receptor		16 p11.2-12.1	GTCTCCAGCATGGGGCAG	AGGAAGAGGTGAAGTCTTGA
X MC1R	Melanocortin-1 Receptor		16 q24	AACCTGCACCTCACCATGTA	TTCCTCTTTTGTTCACATTCAA
BRCA1	Breast Cancer 1		17 q21	CTGTGTGAGAGAAAGAATGG	ACCCTCGTCAGCTTCACCAC
TK1	Thymidine Kinase 1		17 q23.2-q25.3	GCACTGGATGGACCTTCCA	CATGACGATCCAGCATTT
X RCV1	Recoverin		17	TGGTACCAGTCTCTCTGA	AACAATGGATTCCATTAAGA
FECH	Ferrochelatase		18 q21.3	ATAGCATTTACAGTGACCA	CCCACTATCAGGAAGATGA
PRNP	Prion Protein		20 pter-p12	ACCACCAAGGGGAGACCT	TTGGATTTCTTTTTCAGTG
AGT1	Agouti Homologue		20 q11.2	GGATGTCAACCCCTACT	ATGGTCTCATCTCCAGCCA
ADA	Adenosine Deaminase		20 q13.11	ATCCTGTGCTGCATGCGCC	CCAGGTCTCCAAACATGCT
SOD1	Superoxide Dismutase 1		21 q22.1	TGTACCAAGTCGAGGTCTCA	AGGGTCACCGTGGGCTCTT
PDGFB	Platelet Derived Growth Factor B		22 q12.3-q13.1	CGCTGCTCCGGCTGCTGTA	

<sup>1</sup> An X in this column indicates that the PCR product has been confirmed to be the correct locus in the canine genome by sequence analysis

Table 2  
Primer Sets for eleven preferred UM-STSS.

Locus	Index	Species <sup>a</sup>	Primer 1 <sup>b</sup>	Primer 2	Primer 1 Name <sup>c</sup>	Primer 2 Name	P1 A.A. <sup>d</sup>	P2 A.A.	Human Chromosome
CFTR		Human	CTAAGCCATGGCCACAAGCA	CATTGCTTCTATCCTGTGTTG	HCFTREX22D	HCFTREX23U	1346	1407	7q31-q32
			.....T.....	...C.....					
COL10A1		Human	ATTCTCTCCAAAGCTTACCC	GCCACTAGGAATCCTGAGAA	HCOL10A1EX2D	HCOL10A1EX2U	505	693	6q21-q22
		Mouse	.....T.....	.....					
CSF1R		Human	TTCCAAACACAGGGACCTA	CATCCAGGGCGAGAGGA	HCSF1REX3D	HCSF1REX4U	76	161	5q33-q35
		FeSV	.....	.G.....					
CYP11A1		Dog	TTGGACCTCTTTGGAGCTGG	TGGTTGATCTGCCACTGGTT	DCYP11A1EX3D	DCYP11A1EX5U	319	417	7q31
		Human	.....	.....					
DCN1		Human	GTGTATGACAGCTAGCCTGAA	AAGTGAAGCTCCCTCAGATG	HDCN1EX6D	HDCN1EX7U	207	247	12q21-q23
		Rat	.....C.....	.....G.....					
FES		Human	GGGGAACCTTTGGCGAAGTGT	TCCATGACGATGTAGATGGG	HFSEEX14D	HFSEEX15U	573	641	15q25-qter
		FeSV	.....A.....	.....					
GHR		Human	CCAGTTCACAGTTCMAAGAT	TGATTCTTCTGGTCAAGGCA	HGHRX9D	HGHRX10U	301	432	5p13.1-p12
		Rat	.....A.....	A.....					
GLB1		Human	GAATTCATACTGGCTGGCT	CATTCCAATAGGCAAAATTGGT	HGLB1EX8D	HGLB1EX9U	268	319	3pter-p21
		Mouse	..G.....	.....					
PKLR		Human	CGCTCAAGGAGATGATCBA	ATGAGCCGTCGTCAATGTA	HPKLRX4D	HPKLRX6U	72	193	1q21
		Rat	.....	.....					
PVALB		Human	ATGTGAGAGAGTGTTCACAT	TCTTTGTCTCCAGCAGCCAT	HPVALBEX3D	HPVALBEX4U	43	93	22q12-q13.1
		Rat	.....C.....	..C.....					
RB1		Human	GTTCCAGAAATAATCAGATGGT	ACTCATTTCTCCAGTTTCTG	HRB1EX25D	HRB1EX26U	844	905	13q14.2
		Mouse	.....C.....	.....					

<sup>a</sup> Primary index species on top

<sup>b</sup> Dots indicate identical nucleotides

<sup>c</sup> First letter, primary index species; next letters; EX + number = exon number; D = down, U = up

<sup>d</sup> Amino acid over which 5' nucleotide of primer lies



Table 3

Amplification conditions for eleven preferred canine UM-STSs.

Locus	Temperatures (°C)	Times (min)	Size of PCR Product (bp)	
			Human	Dog
CFTR	95, 57, 72	0.5, 1.5, 4	700	1000
COL10A1	94, 57, 72 (hs) <sup>a</sup>	1, 2, 3	384	384
CSF1R	94, 59, 72	1, 2, 3	730	730
CYP1A1	95, 57, 72	0.5, 1.5, 4	700	600
DCN1	94, 57, 72	1, 2, 3	1422	2000
FES	94, 57, 72	0.5, 1, 1.5	484	500
GHR	94, 57, 72	1, 2, 3	765	800
GLB1	94, 57, 72	1, 2, 3	238	240
PKLR	94, 59, 72	1, 2, 3	600	630
PVALB	94, 57, 72 (hs)	0.5, 1.5, 4	1400	1300
RB1	94, 59, 72	1, 2, 3	695	1300

<sup>a</sup> hs indicates 'hot start' used.

-16-

Table 4

Locus	Human	Macaque	Dog	Goat	Pig	Horse	Mouse	Rat
CFTR	<b>+</b> <sup>a</sup>	+	+	-	+	+	+	+
COL10A1	+	+	+	-	+	+	+	+
CSF1R	+	+	+	+	-	-	+	+
CYP1A1	+	+	+	+	+	+	+	+
DCN1	+	+	+	-	+	+	+	+
FES	+	+	+	+	+	+	+	+
GHR	+	+	+	+	+	+	+	+
GLB1	+	+	+	+	+	+	+	+
PKLR	+	+	+	+	+	-	+	+
PVALB	+	+	+	+	+	+	-	-
RB1	+	+	+	-	+	+	+	+

+, amplification; -, no amplification

<sup>a</sup> Bold symbols indicate index

- 17 -

The foregoing discussion discloses and describes merely exemplary embodiments of the present invention. One skilled in the art will readily recognize from such discussion and from the accompanying claims and drawings, that various changes, modifications and variations can be made  
5 therein without departing from the spirit and scope of the invention.

All publications referred to herein are expressly incorporated by reference.

- 18 -

**WE CLAIM:**

1. A primer comprising a polynucleotide, wherein the polynucleotide has a sequence selected from the group consisting of the sequences set forth in Table 1.
- 5 2. A primer comprising a polynucleotide, wherein the polynucleotide has a sequence selected from the group consisting of the sequences set forth in Table 1A.
3. A method for amplifying DNA, comprising the step of performing PCR with the DNA and a primer set selected from the group consisting of the  
10 primer sets of Table 1.
4. A method for amplifying DNA, comprising the step of performing PCR with the DNA and a primer set selected from the group consisting of the primer sets of Table 1A.

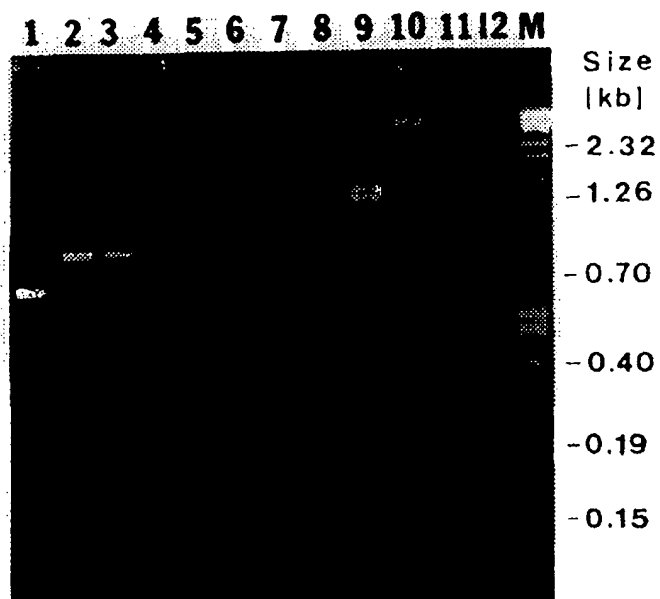


Fig. 1.

2/5

## Figure 2

CFTR	
A.A. 1346	
	intron 22
Dog	- - - - -   I - - - - -
Mouse	- - - - -   I - - - V - - - V - - - - -
Human	E P S A H L D P   V T Y Q I I R R T L K Q A F A
Human	GAACCCAGTGCTCATTGGATCC AGTAACATACCAAATAATTAGAAGAACTCTAAAACAAGCATTGGCT
Mouse	..G.....C...C.A..C.. CA.....G.C...C..C..GT.....C..C...
Dog	G..... ..A.....C
COL10A1	
A.A. 569	
Dog	- - - - - K - - - - - H
Mouse	- I Y E - - - - - S - - - - - K
Human	P F D K I L Y N R Q Q H Y D P R T G I F T C Q
Human	CCATTGATAAAATTTTGTATAACAGGCAACAGCATTATGACCCAAGGACTGGAATCTTTACTTGTTCAG
Mouse	..CA..T..G.G...C...C..T....G.....C.....AT...T.....C...A..
Dog	.....G..C.....A.....A.....C..C..C..C
CSF1R	
A.A. 107	
	intron 3
Dog	- - V - - - Q - - - - - V - G - - - - -
FeLV	- - A - - - Q - - - - - T - L - G - - - - -
Human	D P A R P W N V L A Q E V V V F E D Q D A L L
Human	ACCCTGCCCGGCCCTGGAACGTGCTAGCACAGGAGGTGGTCTGTTCGAGGACCAGGACGCACTACTGTC
FeLV	.....T.....T.....G.....G..C.....A...ACG.....G..A.GT.....T..GT.G....
Dog	.....TT.....T.....G.....G..G.....C.....G.....GG.....T..G..G....
DCN1	
A.A. 231	
	intron 6
Dog	- - - - -   - - - - -
Rat	- - - - - N S - - S   - - - - -
Human	V D A A S L K G L N N L A   K L G L S F N S I S
Human	<u>GTTGATGCAGCTAGCCTGAAAGGACTGAATAATTTGGCTA</u>  AGTTGGGATTGAGTTTCAACAGCATCTCT
Rat	.....C.....A..TC.....T... ..C...T....C.....T.....A.C
Dog	..... .....C.....T.....N
GHR	
A.A. 333	
Dog	- D L - - - - - G - - - - - N - - - - -
Rat	- D A - - - - - - - - - - D - Q - - - - -
Human	D E P D E K T E E S D T D R L L S S D H E K S
Human	TGATGAGCCAGATGAAAAGACTGAGGAATCAGACACAGACAGACTTCTAAGCAGTGACCATGAGAAATCA
Rat	.....TG.G.....G.....A..G.....C.....GA.....G.....
Dog	.....C.T.....C..A.G.....AC.....

3/5

## GLB1

A.A. 268

```

Dog      - - - - - V - - - V - - -
Mouse    - - - - - K - - - V - - K T L - T
Human    E F Y T G W L D H W G Q P H S T I K T E A V A S
Human    GAATTCTATACTGGCTGGCTAGATCACTGGGGCCAACCTCACTCCACAATCAAGACCGAAGCAGTGGCTTCC
Mouse    ..G.....C.....TA...C..T....GG.G..A..TA..A..C....A..
Dog      .....T.....G..A.....A...G.G....T....TC.....

```

## PVALB

A.A. 59

```

Dog      -                               intron 2          - -
Rat      - - - - -                      (bp)              S - - -
Human    I E E D E L G                      F I L K
Human    ATCGAGGAGGATGAGCTGGGgtaagctggagg - 1300 - tttctcctccagATTCATCCTAAAAG
Rat      ..T.....a. - 1500 - .....G.C...T..G..G.
Dog      .....agactcc. - 1300 - .....

```

## RB1

A.A. 890

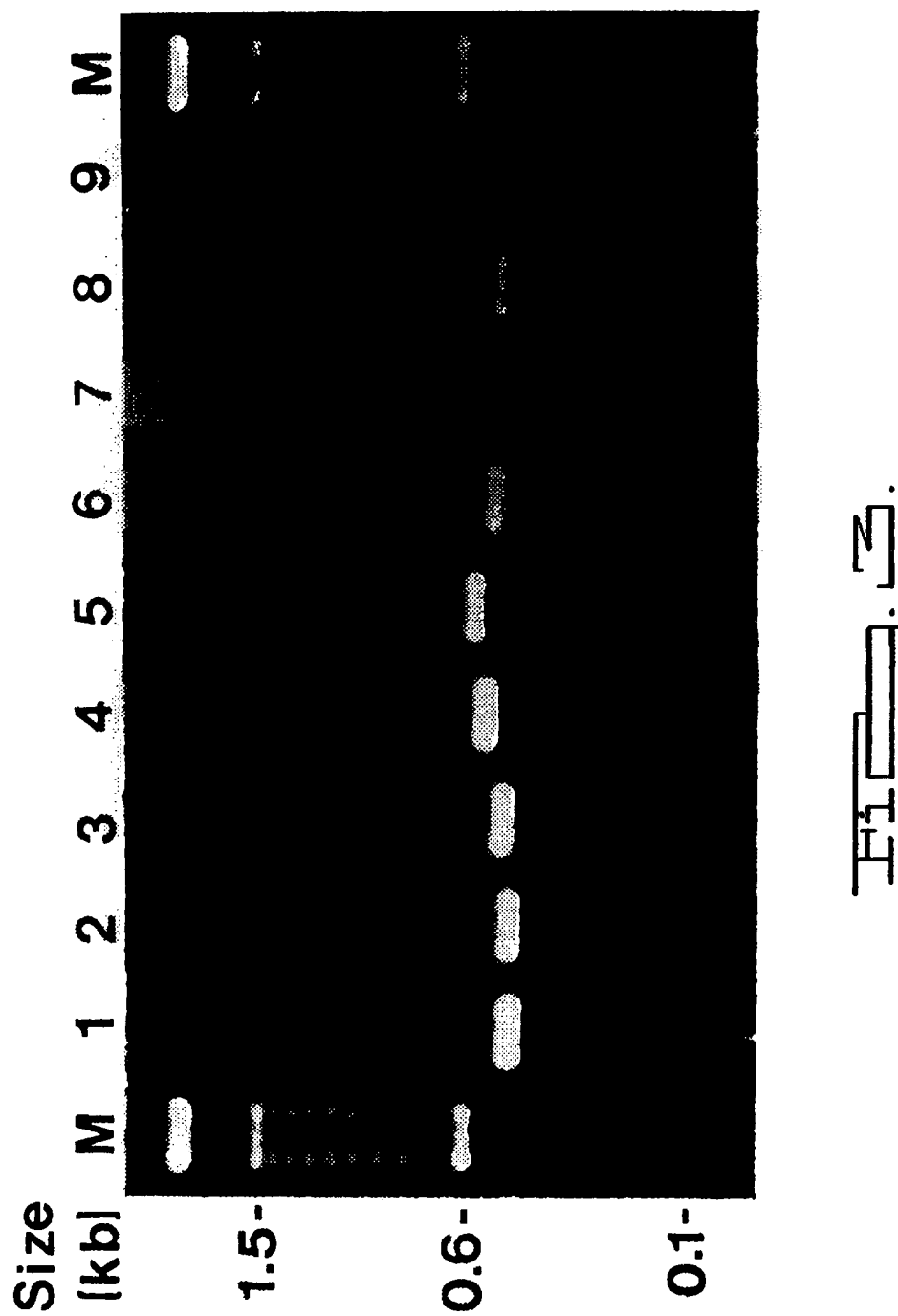
```

Dog      A - - - - - L - - - - - intron 22
Mouse    - G - - - - - N V - - - - - A - - - - -
Human    G S N P P K P L K K L R F D I E G S D E A D G S|
Human    GGAAGCAACCCTCCTAAACCACTGAAAAAAGTACGCTTTGATATTGAAGGATCAGATGAAGCAGATGGAAG|
Mouse    ..CG.....C..C.....CG.G.....C..C..G...G.C.....G..|
Dog      .C.....T.....TG.....C.....

```

Figure 2 (cont.)

4/5





## Figure 4

MAC, CAT, FES I

HUM A D N T L V A V K S C R E T L P P D L K

HUM GCCGACAACACCCTGGTGGCGGTGAAGTCTTGTAGAGAGACGCTCCCACCTGACCTCAAG

MAC .....T.....A.....A.....

CAT .....T.....C.....A.....C.C.....A.....A.....

FES .....T.....C.....A.....C.C.....A.....A.....

DOG .....T..T.....A.....CC.....C.....

COW ..A.....A.....C.C.....A.....

GOA ..A.....A.....C.....A..G..C.....

PIG ..A..T.....A.....CC...A.....

HOR ..T.....A.....CC.....C..G.....

RAT ..A.....C.....T.....C.....N.....

MOU .....T.....C.....NNN.....

HUM A K F L Q E A R

HUM GCCAAGTTTCTACAGGAAGCGAG GTGGGTGATAAACTAATGATCACCACGGGTCCCGCAT

MAC .....C.....G...C...--..CA..A..CT..A...

DOG .....C.....G...C...--..CA..A..CT..A...

CAT .....T.....A..A ..A...AC...AG..C...--..CATAA..T....C

FES .....T.....A..A

COW .....C.....G.AC.CCC...A.TGTA..C...CATA

GOA .....G.....A.....G.AC.CC...A.TGTA..C...T.C.C

PIG .....G.....AG..CC...TGTGATAAAAGA.CC

HOR .....G..A.....C...A...CC...TGGTAT..CTAA.G..

RAT .....G..NNNN.....C.....A.GGGA.CAGT..A..T...TTGTG

MOU .....A.....A....AT

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US97/02403

**A. CLASSIFICATION OF SUBJECT MATTER**

IPC(6) : C07H 21/04; C12P 19/34; C12Q 1/68

US CL : 536/24.31, 24.33; 435/91.2, 6

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 536/24.31, 24.33; 435/91.2, 6

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, DIALOG

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	HUGHES et al. Evolution of the Major Histocompatibility Complex: Independent Origin of Nonclassical Class I Genes in Different Groups of Mammals. Mol. Biol. Evol. 1989, Vol. 6, No. 6, pages 559-579, see Abstract.	1-4
Y	CZELUSNIAK et al. Phylogenetic Origins and Adaptive Evolution of Avian and Mammalian Haemoglobin Genes. Nature. 15 July 1982, Vol. 298, pages 297-300, especially page 300.	1, 3
Y	GRAY, M.W. Mitochondrial Genome Diversity and the Evolution of Mitochondrial DNA. Can. J. Biochem. 1982, Vol. 60, pages 157-171, especially page 165.	1-4

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Z"	document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means		
"P" document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search

12 APRIL 1997

Date of mailing of the international search report

30 APR 1997

Name and mailing address of the ISA/US  
Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

PHUONG BUI

Telephone No. (703) 308-0196

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US97/02403

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	VENTA et al. Organization of the Mouse and Human Carbonic Anhydrase II Genes. Ann. N.Y. Acad. Sci. 1984, Vol. 429, pages 309-323, especially Abstract and page 319.	1, 3
Y	VENTA et al. Comparison of the 5' Regions of Human and Mouse Carbonic Anhydrase II Genes and Identification of Possible Regulatory Elements. Biochim. Biophys. Acta. 1985, Vol. 826, pages 195-201, see entire article.	1, 3
Y	WICHMAN et al. In Search of Retrotransposons: Exploring the Potential of the PCR. Biotechniques. 1992, Vol. 13, No. 2, pages 258-265, see Abstract.	1-4
Y	MAZZARELLA et al. Conserved Sequence-Tagged Sites: A Phylogenetic Approach to Genome Mapping. Proc. Natl. Acad. Sci. USA. May 1992, Vol. 89, pages 3681-3685, see Abstract.	1-4
Y	HINO et al. Universal Mapping Probes and the Origin of Human Chromosome 3. Proc. Natl. Acad. Sci. USA. January 1993, Vol. 90, pages 730-734, see Abstract.	1-4
Y	SKAGGS et al. The Deduced Amino Acid Sequence of Human Carbonic Anhydrase-Related Protein (CARP) is 98% Identical to the Mouse Homologue. Gene. 1993, Vol. 126, pages 291-292, see entire article.	1, 3
Y	MERRYMAN et al. Characterization of a New, Potent, Immunopathogenic Epitope in S-Antigen that Elicits T Cells Expressing V $\beta$ 8 and V $\alpha$ 2-Like Genes. J. Immunol. 01 January 1991, Vol. 146, No. 1, pages 75-80, especially pages 77 and 79.	2, 4
Y	ITAKURA et al. Expression in <i>Escherichia coli</i> of a Chemically Synthesized Gene for the Hormone Somatostatin. Science. 09 December 1977, Vol. 198, No. 4321, pages 1056-1063, see Abstract and pages 1056-1058.	1, 3

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☒ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☒ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**